

On: 04 June 2015, At: 04:08

Publisher: Taylor & Francis

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



[Click for updates](#)

Prion

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/kprn20>

Amyloids or Prions? That is the Question

Raimon Sabate^a, Frederic Rousseau^{bc}, Joost Schymkowitz^{bc}, Cristina Batlle^d & Salvador Ventura^d

^a Departament de Físicoquímica, Facultat de Farmàcia, and Institut de Nanociència i Nanotecnologia (IN²UB), Universitat de Barcelona, Avda. Joan XXIII s/n, E- 08028-Barcelona, Spain

^b VIB Switch Laboratory, VIB, Leuven, Belgium

^c Department for Cellular and Molecular Medicine, KU Leuven, Leuven, Belgium

^d Institut de Biotecnologia i de Biomedicina and Departament de Bioquímica i Biologia Molecular. Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain

Accepted author version posted online: 03 Jun 2015.

To cite this article: Raimon Sabate, Frederic Rousseau, Joost Schymkowitz, Cristina Batlle & Salvador Ventura (2015): Amyloids or Prions? That is the Question, *Prion*, DOI: [10.1080/19336896.2015.1053685](https://doi.org/10.1080/19336896.2015.1053685)

To link to this article: <http://dx.doi.org/10.1080/19336896.2015.1053685>

Disclaimer: This is a version of an unedited manuscript that has been accepted for publication. As a service to authors and researchers we are providing this version of the accepted manuscript (AM). Copyediting, typesetting, and review of the resulting proof will be undertaken on this manuscript before final publication of the Version of Record (VoR). During production and pre-press, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal relate to this version also.

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

Amyloids or Prions? That is the Question.

Raimon Sabate^{1,#}, Frederic Rousseau^{2,3,#}, Joost Schymkowitz^{2,3,#}, Cristina Batlle⁴ and Salvador Ventura^{4,*}

¹Departament de Físicoquímica, Facultat de Farmàcia, and Institut de Nanociència i Nanotecnologia (IN²UB), Universitat de Barcelona, Avda. Joan XXIII s/n, E- 08028-Barcelona, Spain

²VIB Switch Laboratory, VIB, Leuven, Belgium

³Department for Cellular and Molecular Medicine, KU Leuven, Leuven, Belgium

⁴Institut de Biotecnologia i de Biomedicina and Departament de Bioquímica i Biologia Molecular.

Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain

Running title: Determinants of Prion Formation

[#]Equal contribution

*Correspondence: Salvador Ventura, Institut de Biotecnologia i de Biomedicina, Parc de Recerca UAB, Mòdul B, Universitat Autònoma de Barcelona, E-08193 Bellaterra (Barcelona), salvador.ventura@uab.es

ABSTRACT

Despite major efforts devoted to understanding the phenomenon of prion transmissibility, it is still poorly understood how this property is encoded in the amino acid sequence. In recent years, experimental data on yeast prion domains allows to start at least partially decrypting the sequence requirements of prion formation. These experiments illustrate the need for intrinsically disordered sequence regions enriched with a particularly high proportion of glutamine and asparagine. Bioinformatic analysis suggests that these regions strike a balance between sufficient amyloid nucleation propensity on the one hand and disorder on the other, which ensures availability of the amyloid prone regions but entropically prevents unwanted nucleation and facilitates brittleness required for propagation.

Keywords

prions, amyloids, protein intrinsic disorder, Q/N-rich domains, yeast, neurodegenerative diseases.

Abbreviations

Alzheimer's disease, AD; transmissible spongiform encephalopathy, TSE; Parkinson's disease, PD; Creutzfeldt-Jakob disease, CJD; familial amyotrophic lateral sclerosis, fALS; prion forming domain, PFD;

In the cell, proteins attain the native structure through a delicate and balanced network of interactions, where protein folding and aggregation exert as competing pathways^{1,2}. In a protein energy landscape, amyloid-like aggregates represent an energy minimum, being usually thermodynamically more stable than the native conformation. This has led to the hypothesis that the amyloid conformation reflects a universal mode of assembly of polypeptide chains and that native protein conformations are evolutionarily selected metastable states². Amyloids are aggregates displaying fibrillar structure, which is constituted by repetitions of a specific protein in a regular β -sheet conformation that runs perpendicular to the fibril axis³. In humans, amyloids are linked to diseases ranging from neurodegenerative conditions such as Alzheimer's disease (AD), Parkinson's disease (PD) and Creutzfeldt-Jakob disease (CJD), to non-neuronal systemic and localized disorders³. On the other hand functional amyloids, i.e. proteins that exploit the amyloid fold for evolutionarily selected biological functions, have been discovered in diverse species, including human⁴. The roles fulfilled by these functional amyloids range from obligate amyloid structures required for scaffolding and/or movement to conditional amyloids such as the yeast prions that can be triggered by environmental factors⁵. Whether obligate or conditional, the natural selection of amyloid structure as a functional motif indicates that these properties are likely sequence specific. Whereas the attainment and sustainment of the native structure relies on cooperative interactions involving most, if not all, of the sequence of a protein domain^{6,7}, the now widely accepted short stretch hypothesis states that amyloid formation in contrast is nucleated by short regions of the amino acid sequence named aggregation hot-spots (HS), Aggregation Prone Regions (APR) or Aggregation Prone Sequences (APS)^{8,9}. The short stretch model led to the development of over twenty algorithms that more or less successfully predict protein aggregation and amyloid formation based on the identification of specific β -aggregation and amyloid-prone regions in the polypeptide sequences¹⁰⁻¹². In disease-associated amyloids these regions are generally between 5 and 10 residues in length¹³.

Prions are considered a subclass of amyloids in which protein aggregation becomes self-perpetuating and infectious. The phenomenon is known mostly as a neuronal pathology in mammals but in fungi prions play a crucial role in epigenetic inheritance¹⁴⁻¹⁶. Importantly, despite the overlapping conformational properties of amyloids and prions, only a handful of

amyloids are currently considered to display at least partial prion capacity under natural conditions¹⁶. As a result, β -aggregation and amyloid predictors are still a long way from correctly detecting prion sequences in proteomes¹⁷. In fact, the sequence characteristics that make a protein sequence a prion have been elusive for years. Moreover, at first glance, the sequence features conferring prion capacity to prion protein in mammals (PrP) appear to differ remarkably from those determining prion behaviour in fungi.

Yeast prions are the best characterized transmissible amyloids, thus being excellent model systems to address the determinants of concomitant amyloid formation and propagation¹⁸. In these proteins, prion formation from an initially soluble state involves a structural amyloid conversion driven by specific, relatively large, unstructured domains enriched in glutamine/asparagine (Q/N) residues¹⁸. Interestingly, protein domains displaying this sequence signature are over-represented in eukaryotic proteomes relative to prokaryotes, suggesting that prion-like conformational transition might have evolved as a mechanism for regulating gene function at the protein level in eukaryotes¹⁹. It should be mentioned, however, that PrP, the archetypical mammalian prion, lacks these sequential features.

In order to explore the repertoire of prion proteins in *Saccharomyces cerevisiae* the Lindquist group conducted a genome-wide bioinformatics survey using a hidden Markov sequence model to identify putative candidates on the basis of their compositional similarity to known prion forming domains (PFDs)²⁰ and used experimental validation to identify the bone-fide PFDs in their predictions. These results are at the core of several algorithms for prion domain prediction, all relying on the analysis of amino acid sequences^{17, 21-24}. These programs are constructed on two alternative models for amyloid formation by prion-like domains (**Figure 1**): (1) The compositional model relying on the establishment of a large number of weak interactions¹⁷ and (2) our model, which proposes a classical nucleation by short amyloidogenic stretches, whose amyloid propensity is modulated by the structural context²⁴. Despite the mechanistic difference between algorithms, the advent of accurate computational tools to detect yeast prion domains opens new and exciting possibilities, allowing the exploration of proteomes for the discovery of novel and hitherto unexpected Q/N-enriched domains that may drive conformational conversion in novel prion proteins. Indeed, recent studies have revealed that over 250 human proteins display prion-like stretches in regions with high presence of uncharged polar residues and

glycine, including several heterogeneous nuclear ribonucleoproteins (hnRNPs) related to neurodegenerative diseases such as familial amyotrophic lateral sclerosis (fALS) ²⁵.

In the light of these advances, the requirements for a polypeptide sequence to act as a prion begin to be defined. In our view, for a protein sequence to become a Q/N enriched prion, three essential conditions appear to be required:

Requirement 1: A short amyloid-prone region able to trigger amyloid formation in a sequence specific manner. These amyloid cores should, however, possess distinctive features, since both a high aggregation rate and an elevated fragmentation capacity are necessary in prions in order to attain the number of propagons or seeds required for spreading and propagation ^{26,27}. Thus, while a certain amyloid nucleation capacity favouring a sufficiently high aggregation rate is absolutely necessary, the final amyloid aggregate should at the same time display brittleness, a property that facilitates an increase in the number of nucleation events per cell. Accordingly, in contrast to most amyloids, the aggregation reaction should not be nucleated in PFDs by an extremely strong and highly hydrophobic amyloid core.

Requirement 2: The amyloid-prone region has to be located in a structurally disordered region, that permits its self-assembly without the necessity of conformational unfolding. The PFDs of all known Q/N enriched yeast prions display this property ^{18,20}. The location of the amyloid core in large unstructured regions favours the acquisition of the β -cross motif without large conformational rearrangements and may at the same time promote the brittleness mentioned in requirement 1. Moreover, the disordered region may act as a so-called entropic bristle ^{27,28}, which would reduce the overall aggregation propensity and could allow for a better biological control of the nucleation event, which is discussed more in detail in requirement 3.

Requirement 3: PFDs have to possess an amino acid composition allowing the protein to remain in a soluble state under physiological conditions while keeping intact a cryptic amyloid capacity. Stress situations promoting increased local protein concentration, as well as the presence of preformed amyloid seeds, might alter the delicate equilibrium between native/soluble and amyloid/insoluble states, providing means to control the nucleation of amyloid aggregation and hence the onset of the prion phenotype. If we recapitulate requirements (1) and (2), clustering in the same sequence region amino acid residues with a significant amyloid propensity together

with residues promoting structural disorder would favour prion capacity. Only five amino acid residues seem to unite these two essential properties, *i.e.* amyloid propensity and structural disorder, according to FoldIndex²⁹ and Waltz³⁰ algorithms: N, Q, Y, S and W. (**Figure 2**). Interestingly enough N, Q, Y, S are, in this order, the most over-represented residues in bona-fide prion domains, relative to their frequency in the protein universe²², with odds ratios of 5.70, 4.13, 1.72 and 1.66, respectively. In this context, N and Q residues show medium amyloid propensity, allowing the formation of amyloids with moderate strength, while at the same time are the amyloidogenic residues that more benefic disorder. This provides a rational basis for the strong over-representation of these particular residues in the PFDs of yeast. In good agreement with its higher frequency, N is the residue that best balances amyloid and disorder propensity and thus the preferred residue to support prion behaviour. It is important to point out, however, that poly-N or poly-Q, these later being involved in a number of ataxias³¹, are not expected to display a prion-like behaviour since they lack requirement 1, that is, a specific region able to selectively nucleate ordered amyloid formation.

Hydrophobic amino acids are under-represented in Q/N-rich yeast PFDs, likely because a high proportion of these residues would render the protein excessively aggregation-prone and/or result in too strong amyloid assemblies. Despite the presence of a reduced number of hydrophobic residues in PFDs has been shown to bust prion formation and amyloid formation³², Y is the only hydrophobic residue over-represented in these domains. It has been proposed that this might respond to the fact that aromatic residues might facilitate both prion formation and chaperone dependent prion propagation³³. However, F is indeed under-represented in PFDs with an odds ratio of 0.72 and the Y/F relationship between odds ratios in PFDs is 2.4, suggesting that the additional hydroxyl group in Y should provide an additional advantage, which in our opinion is allowing a better balance of amyloid propensity and intrinsic disorder. Despite its aromatic character, W is one of the most under-represented residues in prion domains with an odd ratio of 0.091, only C, which is able to crosslink covalently polypeptide chains, being less frequent²². The absence of W in PFDs is best explained by its particular structure, wherein the indole group may not be easily placed in β -cross structures due to of

steric impediments, being indeed depleted relative to F and Y in functional and pathogenic amyloids³⁴.

The two alternative models used to identify prion domains (**Figure 1**) coincide in the requirement of a relatively large disordered region in yeast PFDs. However, one prion model support the view of amyloid formation in PFDs resulting from a bias in sequence composition favouring a large number of weak interactions over a wide sequence stretch¹⁷ whereas the alternative model supports prion behaviour to emerge from the preferential nucleation by specific and localized short amyloid-prone stretches embedded in the wider disordered region²⁴. Despite the apparent contradiction between these two views, indeed the second model just pursues to delimitate the aggregation driving force of the amyloid-cores embedded in the prions domains defined by the compositional model. In this way, the first composition based methods to predict potential yeast prions proposed a minimal core of 60 residues²⁰. This further evolved into a method employing a 41 amino acid sliding window for compositional analysis, denoting that the initial 60 residues window was larger than actually required¹⁷. We proposed to reduce this size even further to account for a 21 amino acid core, based on the length of the core of HET-s PFD, the unique protein for which an atomic-resolution structure of the infectious fibrillar state available to date^{24, 35}, which displays a β -arcade conformation³⁶. The excellent performance of our method, based on a preferential nucleation by a short amyloid-prone stretches, lead us to believe not only that a 21 residues core is indeed sufficient for prediction, but also that the classical short stretch nucleation model applies to prions in a similar manner as it does for classical amyloids, the main differences being that in prions amyloid nucleating stretches might fold into β -strand-turn- β -strand elements and that they are strongly modulated by the entropy of the sequence context in which they are embedded, i.e. the degree of structural disorder will determine both the sensitivity for amyloid nucleation as well as the ability of formed fibrils to break up and provide additional propagons.

ACKNOWLEDGMENTS

We thank Pol Ventura for his comments on Figure 2. This work was supported in part by Ministerio de Economía y Competividad, Spain [BFU2013-44763 to S.V.]; by SUDOE, INTERREG IV B, FEDER [SOE4/P1/E831to S.V.]; by ICREA [ICREA Academia 2009 to S.V.];

by the Ramón y Cajal Programme from Ministerio de Ciencia e Innovación [RYC-2011-07987 to R.S.]. The Switch Laboratory was supported by grants from VIB, University of Leuven, the Funds for Scientific Research Flanders (FWO), the Flanders Institute for Science and Technology (IWT) and the Federal Office for Scientific Affairs of Belgium (Belspo, IAP network P7/16).

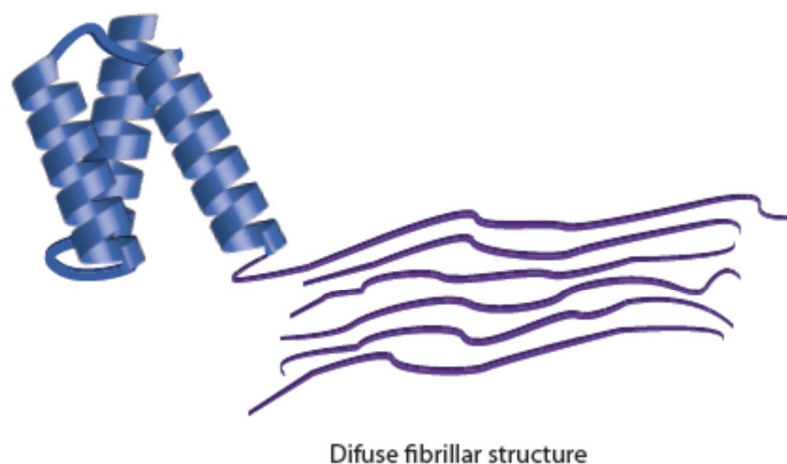
Accepted Manuscript

REFERENCES

1. Jahn TR, Radford SE. The Yin and Yang of protein folding. *Febs J* 2005; 272:5962-70.
2. Jahn TR, Radford SE. Folding versus aggregation: polypeptide conformations on competing pathways. *Arch Biochem Biophys* 2008; 469:100-17.
3. Chiti F, Dobson CM. Protein misfolding, functional amyloid, and human disease. *Annu Rev Biochem* 2006; 75:333-66.
4. Fowler DM, Koulov AV, Balch WE, Kelly JW. Functional amyloid--from bacteria to humans. *Trends in biochemical sciences* 2007; 32:217-24.
5. Falsone A, Falsone SF. Legal but lethal: functional protein aggregation at the verge of toxicity. *Frontiers in cellular neuroscience* 2015; 9:45.
6. Levinthal C. Are there pathways for protein folding? . *J Chim Phys* 1968; 65:44-5.
7. Zwanzig R, Szabo A, Bagchi B. Levinthal's paradox. *Proc Natl Acad Sci U S A* 1992; 89:20-2.
8. Lopez de la Paz M, Serrano L. Sequence determinants of amyloid fibril formation. *Proc Natl Acad Sci U S A* 2004; 101:87-92.
9. Sanchez de Groot N, Pallares I, Aviles FX, Vendrell J, Ventura S. Prediction of "hot spots" of aggregation in disease-linked polypeptides. *BMC Struct Biol* 2005; 5:18.
10. Belli M, Ramazzotti M, Chiti F. Prediction of amyloid aggregation in vivo. *EMBO Rep* 2011; 12:657-63.
11. Castillo V, Grana-Montes R, Sabate R, Ventura S. Prediction of the aggregation propensity of proteins from the primary sequence: aggregation properties of proteomes. *Biotechnol J* 2011; 6:674-85.
12. Espargaro A, Busquets MA, Estelrich J, Sabate R. Predicting the aggregation propensity of prion sequences. *Virus Res* 2015.
13. Rousseau F, Serrano L, Schymkowitz JW. How evolutionary pressure against protein aggregation shaped chaperone specificity. *Journal of molecular biology* 2006; 355:1037-47.
14. Aguzzi A, Calella AM. Prions: protein aggregation and infectious diseases. *Physiol Rev* 2009; 89:1105-52.
15. Chien P, Weissman JS, DePace AH. Emerging principles of conformation-based prion inheritance. *Annu Rev Biochem* 2004; 73:617-56.
16. Sabate R. When amyloids become prions. *Prion* 2014; 8.
17. Toombs JA, Petri M, Paul KR, Kan GY, Ben-Hur A, Ross ED. De novo design of synthetic prion domains. *Proc Natl Acad Sci U S A* 2012; 109:6519-24.
18. Uptain SM, Lindquist S. Prions as protein-based genetic elements. *Annu Rev Microbiol* 2002; 56:703-41.
19. Michelitsch MD, Weissman JS. A census of glutamine/asparagine-rich regions: implications for their conserved function and the prediction of novel prions. *Proc Natl Acad Sci U S A* 2000; 97:11910-5.
20. Alberti S, Halfmann R, King O, Kapila A, Lindquist S. A systematic survey identifies prions and illuminates sequence features of prionogenic proteins. *Cell* 2009; 137:146-58.
21. Espinosa Angarica V, Angulo A, Giner A, Losilla G, Ventura S, Sancho J. PrionScan: an online database of predicted prion domains in complete proteomes. *BMC Genomics* 2014; 15:102.

22. Espinosa Angarica V, Ventura S, Sancho J. Discovering putative prion sequences in complete proteomes using probabilistic representations of Q/N-rich domains. *BMC Genomics* 2013; 14:316.
23. Lancaster AK, Nutter-Upham A, Lindquist S, King OD. PLAAC: a web and command-line application to identify proteins with prion-like amino acid composition. *Bioinformatics* 2014.
24. Sabate R, Rousseau F, Schymkowitz J, Ventura S. What makes a protein sequence a prion? *PLoS Comput Biol* 2015; 11:e1004013.
25. Kim HJ, Kim NC, Wang YD, Scarborough EA, Moore J, Diaz Z, et al. Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature* 2013; 495:467-73.
26. Tanaka M, Collins SR, Toyama BH, Weissman JS. The physical basis of how prion conformations determine strain phenotypes. *Nature* 2006; 442:585-9.
27. Derdowski A, Sindi SS, Klaips CL, DiSalvo S, Serio TR. A size threshold limits prion transmission and establishes phenotypic diversity. *Science* 2010; 330:680-3.
28. Grana-Montes R, Marinelli P, Reverter D, Ventura S. N-terminal protein tails act as aggregation protective entropic bristles: the SUMO case. *Biomacromolecules* 2014; 15:1194-203.
29. Prilusky J, Felder CE, Zeev-Ben-Mordehai T, Rydberg EH, Man O, Beckmann JS, et al. FoldIndex: a simple tool to predict whether a given protein sequence is intrinsically unfolded. *Bioinformatics* 2005; 21:3435-8.
30. Maurer-Stroh S, Debulpaep M, Kuemmerer N, Lopez de la Paz M, Martins IC, Reumers J, et al. Exploring the sequence determinants of amyloid structure using position-specific scoring matrices. *Nat Meth* 2010; 7:237-42.
31. Paulson HL, Perez MK, Trottier Y, Trojanowski JQ, Subramony SH, Das SS, et al. Intranuclear inclusions of expanded polyglutamine protein in spinocerebellar ataxia type 3. *Neuron* 1997; 19:333-44.
32. Ross ED, Toombs JA. The effects of amino acid composition on yeast prion formation and prion domain interactions. *Prion* 2010; 4:60-5.
33. Gonzalez Nelson AC, Paul KR, Petri M, Flores N, Rogge RA, Cascarina SM, et al. Increasing prion propensity by hydrophobic insertion. *PLoS One* 2014; 9:e89286.
34. Gazit E. A possible role for pi-stacking in the self-assembly of amyloid fibrils. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 2002; 16:77-83.
35. Wasmer C, Lange A, Van Melckebeke H, Siemer AB, Riek R, Meier BH. Amyloid fibrils of the HET-s(218-289) prion form a beta solenoid with a triangular hydrophobic core. *Science* 2008; 319:1523-6.
36. Kajava AV, Baxa U, Steven AC. Beta arcades: recurring motifs in naturally occurring and disease-related amyloid fibrils. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 2010; 24:1311-9.

COMPOSITIONAL MODEL



AMYLOID-STRETCH MODEL

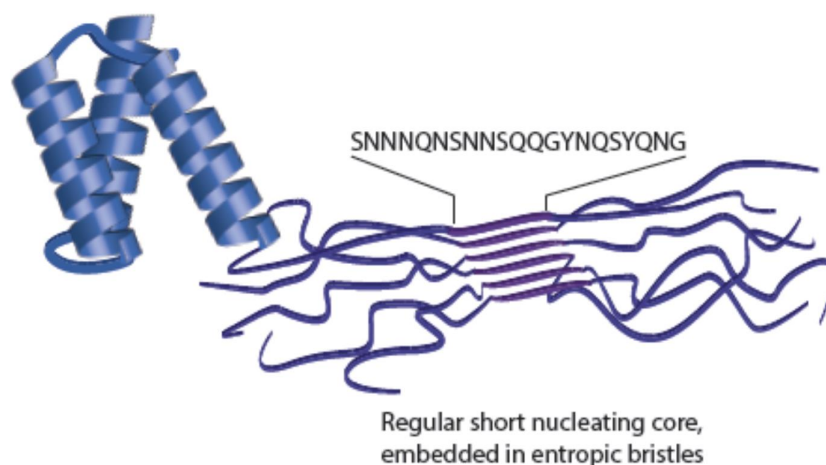


Figure 1. Two models for amyloid structure formation in Q/N-rich prion-like domains. The compositional model relies on the establishment a large number of weak interactions whereas the amyloid-stretch model proposes the existence of a preferential short nucleating sequence whose amyloid propensity is modulated by its structural context.

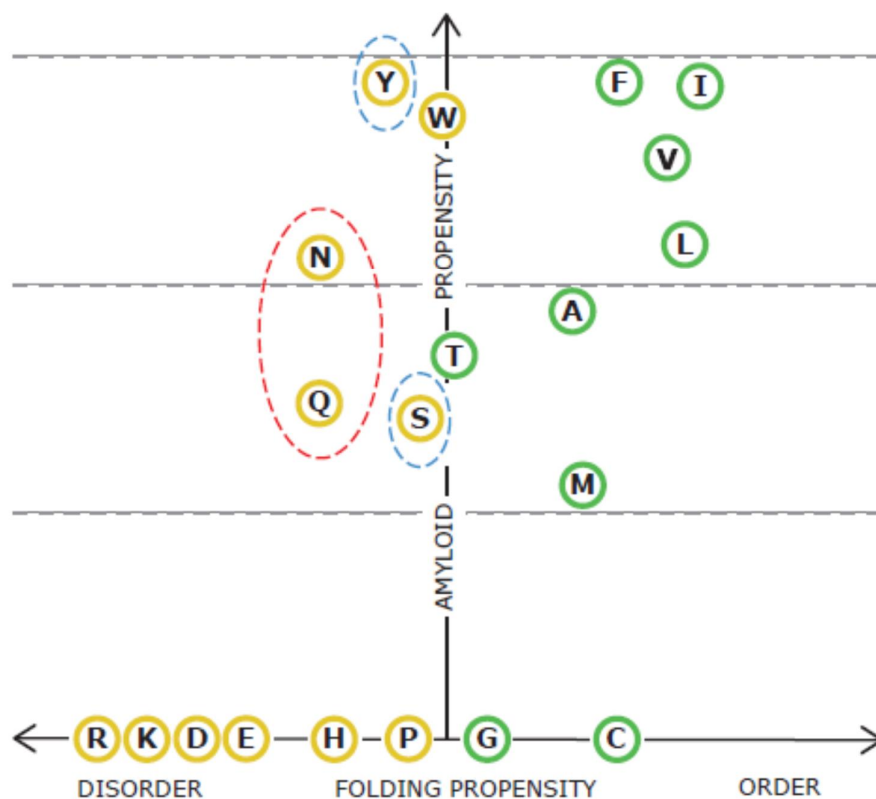


Figure 2. Balance between amyloid and structural propensities in natural amino acids.

Residues rendering ordered and disordered 21-residues long homo-polymers according to FoldIndex¹⁴ are shown in green and yellow circles, respectively. The amyloid propensity of these stretches was calculated with Waltz³⁰. The four more over-represented residues in yeast PFDs are circled by discontinuous lines, red indicates odd ratios > 4.0 and blue odd ratios > 1.5, relative to the composition of the protein universe²².